



1/36

SEQUENCE LISTING

<110> Gray, John T.
Mulligan, Richard C.

<120> Packaging Cell Lines

<130> CMCC693p2A

<140> US 09/393,795

<141> 1999-09-10

<150> US 60/100,063

<151> 1998-09-12

<150> US 60/100,022

<151> 1998-09-11

<160> 12

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 396

<212> DNA

<213> Homo sapien

<400> 1

atgggtgcga gagcgctcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 60
ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcagggag 120
ctagaacgat tcgcagttaa tcctggcctg ttagaaacat cagaaggctg tagacaaata 180
ctgggacagc tacaaccatc cttcagaca ggatcagaag aacttagatc attatataat 240
acagtagcaa ccctctattg tgtgcatcaa aggatagaga taaaagacac caaggaagct 300
ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 360
gacacaggac acagcaatca ggtcagccaa aattac 396

<210> 2

<211> 1503

<212> DNA

<213> Homo sapien

<220>

<221> CDS

<222> (1) ... (1503)

<400> 2

atg ggt gcg aga gcg tcg gta tta agc ggg gga gaa tta gat aaa tgg 48
Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Lys Trp
1 5 10 15

gaa aaa att cgg tta agg cca ggg gga aag aaa caa tat aaa cta aaa	96
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Gln Tyr Lys Leu Lys	
20 25 30	
cat ata gta tgg gca agc agg gag cta gaa cga ttc gca gtt aat cct	144
His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro	
35 40 45	
ggc ctt tta gag aca tca gaa ggc tgt aga caa ata ctg gga cag cta	192
Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu	
50 55 60	
caa cca tcc ctt cag aca gga tca gaa gaa ctt aga tca tta tat aat	240
Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn	
65 70 75 80	
aca ata gca gtc ctc tat tgt gtg cat caa agg ata gat gta aaa gac	288
Thr Ile Ala Val Leu Tyr Cys Val His Gln Arg Ile Asp Val Lys Asp	
85 90 95	
acc aag gaa gcc tta gat aag ata gag gaa gag caa aac aaa agt aag	336
Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys	
100 105 110	
aaa aag gca cag caa gca gca gct gac aca gga aac aac agc cag gtc	384
Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Asn Ser Gln Val	
115 120 125	
agc caa aat tac cct ata gtg cag aac ctc cag ggg caa atg gta cat	432
Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His	
130 135 140	
cag gcc ata tca cct aga act tta aat gca tgg gta aaa gta gta gaa	480
Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu	
145 150 155 160	
gag aag gct ttc agc cca gaa gta ata ccc atg ttt tca gca tta tca	528
Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser	
165 170 175	
gaa gga gcc acc cca caa gat tta aat acc atg cta aac aca gtg ggg	576
Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly	
180 185 190	
gga cat caa gca gcc atg caa atg tta aaa gag acc atc aat gag gaa	624
Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu	
195 200 205	
gct gca gaa tgg gat aga ttg cat cca gtg cat gca ggg cct att gca	672
Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala	
210 215 220	

cca ggc cag atg aga gaa cca agg gga agt gac ata gca gga act act	720
Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr	
225 230 235 240	
agt acc ctt cag gaa caa ata gga tgg atg aca cat aat cca cct atc	768
Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr His Asn Pro Pro Ile	
245 250 255	
cca gta gga gaa atc tat aaa aga tgg ata atc ctg gga tta aat aaa	816
Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys	
260 265 270	
ata gta aga atg tat agc cct acc agc att ctg gac ata aga caa gga	864
Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly	
275 280 285	
cca aag gaa ccc ttt aga gac tat gta gac cga ttc tat aaa act cta	912
Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu	
290 295 300	
aga gcc gag caa gct tca caa gag gta aaa aat tgg atg aca gaa acc	960
Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr	
305 310 315 320	
ttg ttg gtc caa aat gcg aac cca gat tgt aag act att tta aaa gca	1008
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala	
325 330 335	
ttg gga cca gga gcg aca cta gaa gaa atg atg aca gca tgt cag gga	1056
Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly	
340 345 350	
gtg ggg gga ccc ggc cat aaa gca aga gtt ttg gct gaa gca atg agc	1104
Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser	
355 360 365	
caa gta aca aat cca gct acc ata atg ata cag aaa ggc aat ttt agg	1152
Gln Val Thr Asn Pro Ala Thr Ile Met Ile Gln Lys Gly Asn Phe Arg	
370 375 380	
aac caa aga aag act gtt aag tgt ttc aat tgt ggc aaa gaa ggg cac	1200
Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His	
385 390 395 400	
ata gcc aaa aat tgc agg gcc cct agg aaa aag ggc tgt tgg aaa tgt	1248
Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys	
405 410 415	
gga aag gaa gga cac caa atg aaa gat tgt act gag aga cag gct aat	1296
Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn	
420 425 430	

ttt tta ggg aag atc tgg cct tcc cac aag gga agg cca ggg aat ttt	1344
Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe	
435 440 445	

ctt cag agc aga cca gag cca aca gcc cca cca gaa gag agc ttc agg	1392
Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg	
450 455 460	

ttt ggg gaa gag aca aca act ccc tct cag aag cag gag ccg ata gac	1440
Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp	
465 470 475 480	

aag gaa ctg tat cct tta gct tcc ctc aga tca ctc ttt ggc agc gac	1488
Lys Glu Leu Tyr Pro Leu Ala Ser Leu Arg Ser Leu Phe Gly Ser Asp	
485 490 495	

ccc tcg tca caa taa	1503
Pro Ser Ser Gln *	
500	

<210> 3
 <211> 500
 <212> PRT
 <213> Homo sapien

<400> 3	
Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Lys Trp	
1 5 10 15	
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Gln Tyr Lys Leu Lys	
20 25 30	
His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro	
35 40 45	
Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu	
50 55 60	
Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn	
65 70 75 80	
Thr Ile Ala Val Leu Tyr Cys Val His Gln Arg Ile Asp Val Lys Asp	
85 90 95	
Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys	
100 105 110	
Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Asn Ser Gln Val	
115 120 125	
Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His	
130 135 140	
Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu	
145 150 155 160	
Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser	
165 170 175	
Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly	
180 185 190	

Gly	His	Gln	Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu		
		195					200					205					
Ala	Ala	Glu	Trp	Asp	Arg	Leu	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala		
		210					215				220						
Pro	Gly	Gln	Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr		
225					230					235					240		
Ser	Thr	Leu	Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	His	Asn	Pro	Pro	Ile		
				245					250					255			
Pro	Val	Gly	Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys		
			260					265					270				
Ile	Val	Arg	Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly		
		275					280					285					
Pro	Lys	Glu	Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu		
		290				295					300						
Arg	Ala	Glu	Gln	Ala	Ser	Gln	Glu	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr		
305					310					315					320		
Leu	Leu	Val	Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala		
				325					330					335			
Leu	Gly	Pro	Gly	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly		
			340					345					350				
Val	Gly	Gly	Pro	Gly	His	Lys	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met	Ser		
		355					360					365					
Gln	Val	Thr	Asn	Pro	Ala	Thr	Ile	Met	Ile	Gln	Lys	Gly	Asn	Phe	Arg		
		370				375					380						
Asn	Gln	Arg	Lys	Thr	Val	Lys	Cys	Phe	Asn	Cys	Gly	Lys	Glu	Gly	His		
385					390					395					400		
Ile	Ala	Lys	Asn	Cys	Arg	Ala	Pro	Arg	Lys	Lys	Gly	Cys	Trp	Lys	Cys		
				405					410					415			
Gly	Lys	Glu	Gly	His	Gln	Met	Lys	Asp	Cys	Thr	Glu	Arg	Gln	Ala	Asn		
			420					425					430				
Phe	Leu	Gly	Lys	Ile	Trp	Pro	Ser	His	Lys	Gly	Arg	Pro	Gly	Asn	Phe		
		435				440						445					
Leu	Gln	Ser	Arg	Pro	Glu	Pro	Thr	Ala	Pro	Pro	Glu	Glu	Ser	Phe	Arg		
		450				455					460						
Phe	Gly	Glu	Glu	Thr	Thr	Thr	Pro	Ser	Gln	Lys	Gln	Glu	Pro	Ile	Asp		
465					470					475					480		
Lys	Glu	Leu	Tyr	Pro	Leu	Ala	Ser	Leu	Arg	Ser	Leu	Phe	Gly	Ser	Asp		
				485					490					495			
Pro	Ser	Ser	Gln														
			500														

<210> 4

<211> 1503

<212> DNA

<213> Artificial Sequence

<220>

<223> Codon optimized form of HIV gag coding region

<221> CDS

<222> (1)...(1503)

<400> 4

atg ggc gcc cgc gcc tcc gtg ctg tcc ggc ggc gag ctg gac aag tgg	48
Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Lys Trp	
1 5 10 15	
 gag aag atc cgc ctg cgc ccc ggc ggc aag aag cag tac aag ctg aag	96
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Gln Tyr Lys Leu Lys	
20 25 30	
 cac atc gtg tgg gcc tcc cgc gag ctg gag cgc ttc gcc gtg aac ccc	144
His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro	
35 40 45	
 ggc ctg ctg gag acc tcc gag ggc tgc cgc cag atc ctg ggc cag ctg	192
Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu	
50 55 60	
 cag ccc tcc ctg caa acc ggc tcc gag gag ctg cgc tcc ctg tac aac	240
Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn	
65 70 75 80	
 acc atc gcc gtg ctg tac tgc gtg cac cag cgc atc gac gtg aag gac	288
Thr Ile Ala Val Leu Tyr Cys Val His Gln Arg Ile Asp Val Lys Asp	
85 90 95	
 acc aag gag gcc ctg gac aag atc gag gag gag cag aac aag tcc aag	336
Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys	
100 105 110	
 aag aag gcc cag cag gcc gcc gcc gac acc ggc aac aac tcc cag gtg	384
Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Asn Ser Gln Val	
115 120 125	
 tcc cag aac tac ccc atc gtg cag aac ctg cag ggc cag atg gtg cac	432
Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His	
130 135 140	
 cag gcc atc tcc ccc cgc acc ctg aac gcc tgg gtg aag gtg gtg gag	480
Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu	
145 150 155 160	
 gag aag gcc ttc tcc ccc gaa gtc atc ccc atg ttc tcc gcc ctg tcc	528
Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser	
165 170 175	
 gag ggc gcc acc ccc cag gac ctg aac acc atg ctg aac acc gtg ggc	576
Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly	
180 185 190	

ggc cac cag gcc gcc atg cag atg ctg aag gag acc atc aac gag gag Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu 195 200 205	624
gcc gcc gag tgg gac cgc ctg cac ccc gtg cac gcc ggc ccc atc gcc Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala 210 215 220	672
ccc ggc cag atg cgc gag ccc cgc ggc tcc gac atc gcc ggc acc acc Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr 225 230 235 240	720
tcc acc ctg caa gag cag atc ggc tgg atg acc cac aac ccc ccc atc Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr His Asn Pro Pro Ile 245 250 255	768
ccc gtg ggc gag atc tac aag cgc tgg atc atc ctg ggc ctg aac aag Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys 260 265 270	816
atc gtg cgc atg tac tcc ccc acc tcc atc ctg gac atc cgc cag ggc Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly 275 280 285	864
ccc aag gag ccc ttc cgc gac tac gtg gac cgc ttc tac aag acc ctg Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu 290 295 300	912
cgc gcc gag cag gcc tcc cag gag gta aag aac tgg atg acc gag acc Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr 305 310 315 320	960
ctg ctg gtg cag aac gcc aac ccc gac tgc aag acc atc ctg aag gcc Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala 325 330 335	1008
ctg ggc ccc ggc gcc acc ctg gag gag atg atg acc gcc tgc cag ggc Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly 340 345 350	1056
gtg ggc ggc ccc ggc cac aag gcc cgc gtg ctg gcc gag gcc atg tcc Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser 355 360 365	1104
caa gtc acc aac ccc gcc acc atc atg atc cag aag ggc aac ttc cgc Gln Val Thr Asn Pro Ala Thr Ile Met Ile Gln Lys Gly Asn Phe Arg 370 375 380	1152
aac cag cgc aag acc gtg aag tgc ttc aac tgc ggc aag gag ggc cac Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His 385 390 395 400	1200

```

atc gcc aag aac tgc cgc gcc ccc cgc aag aag ggc tgc tgg aag tgc      1248
Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
      405                      410                      415

ggc aag gag ggc cac cag atg aaa gat tgt act gag aga cag gct aat      1296
Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
      420                      425                      430

ttt tta ggg aag atc tgg cct tcc cac aag gga agg cca ggg aat ttt      1344
Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe
      435                      440                      445

ctt cag agc aga cca gag cca aca gcc cca cca gaa gag agc ttc agg      1392
Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
      450                      455                      460

ttt ggg gaa gag aca aca act ccc tct cag aag cag gag ccg ata gac      1440
Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
465                      470                      475                      480

aag gaa ctg tat cct tta gct tcc ctc aga tca ctc ttt ggc agc gac      1488
Lys Glu Leu Tyr Pro Leu Ala Ser Leu Arg Ser Leu Phe Gly Ser Asp
      485                      490                      495

ccc tcg tca caa taa      1503
Pro Ser Ser Gln *
      500

```

```

<210> 5
<211> 500
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Codon optimized form of HIV gag coding region

```

```

<400> 5
Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Lys Trp
 1      5      10      15
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Gln Tyr Lys Leu Lys
 20      25      30
His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 35      40      45
Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 50      55      60
Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 65      70      75      80
Thr Ile Ala Val Leu Tyr Cys Val His Gln Arg Ile Asp Val Lys Asp
 85      90      95
Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
100      105      110

```


Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Asn Ser Gln Val
 115 120 125
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His
 130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr His Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 325 330 335
 Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Thr Asn Pro Ala Thr Ile Met Ile Gln Lys Gly Asn Phe Arg
 370 375 380
 Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His
 385 390 395 400
 Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
 405 410 415
 Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
 420 425 430
 Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe
 435 440 445
 Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
 450 455 460
 Phe Gly Glu Glu Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
 465 470 475 480
 Lys Glu Leu Tyr Pro Leu Ala Ser Leu Arg Ser Leu Phe Gly Ser Asp
 485 490 495
 Pro Ser Ser Gln
 500

<210> 6
 <211> 3012
 <212> DNA
 <213> Homo sapien

<220>
 <221> CDS
 <222> (1)...(3012)

<400> 6

ttt ttt agg gaa gat ctg gcc ttc cca caa ggg aag gcc agg gaa ttt	48
Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe	
1 5 10 15	
tct tca gag cag acc aga gcc aac agc ccc acc aga aga gag ctt cag	96
Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln	
20 25 30	
gtt tgg gga aga gac aac aac tcc ctc tca gaa gca gga gcc gat aga	144
Val Trp Gly Arg Asp Asn Asn Ser Leu Ser Glu Ala Gly Ala Asp Arg	
35 40 45	
caa gga act gta tcc ttt agc ttc cct cag atc act ctt tgg cag cga	192
Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg	
50 55 60	
ccc ctc gtc aca ata aag ata ggg ggg caa tta aag gaa gct cta tta	240
Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu	
65 70 75 80	
gat aca gga gca gat gat aca gta tta gaa gaa atg aat ttg cca gga	288
Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Met Asn Leu Pro Gly	
85 90 95	
aga tgg aaa cca aaa atg ata ggg gga att gga ggt ttt atc aaa gta	336
Arg Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val	
100 105 110	
gga cag tat gat cag ata ctc ata gaa atc tgc gga cat aaa gct ata	384
Gly Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile	
115 120 125	
ggg aca gta tta gta gga cct aca cct gtc aac ata att gga aga aat	432
Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn	
130 135 140	
ctg ttg act cag att ggc tgc act tta aat ttt ccc att agt cct att	480
Leu Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile	
145 150 155 160	

gag act gta cca gta aaa tta aag cca gga atg gat ggc cca aaa gtt Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val 165 170 175	528
aaa caa tgg cca ttg aca gaa gaa aaa ata aaa gca tta gta gaa att Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile 180 185 190	576
tgt aca gaa atg gaa aag gaa gga aaa att tca aaa att ggg cct gaa Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu 195 200 205	624
aat cca tac aat act cca gta ttt gcc ata aag aaa aaa gac agt act Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr 210 215 220	672
aaa tgg aga aaa tta gta gat ttc aga gaa ctt aat aag aga act caa Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln 225 230 235 240	720
gat ttc tgg gaa gtt caa tta gga ata cca cat cct gca ggg tta aaa Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys 245 250 255	768
cag aaa aaa tca gta aca gta ctg gat gtg ggc gat gca tat ttt tca Gln Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser 260 265 270	816
gtt ccc tta gat aaa gac ttc agg aag tat act gca ttt acc ata cct Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro 275 280 285	864
agt ata aac aat gag aca cca ggg att aga tat cag tac aat gtg ctt Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu 290 295 300	912
cca cag gga tgg aaa gga tca cca gca ata ttc cag tgt agc atg aca Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr 305 310 315 320	960
aaa atc tta gag cct ttt aga aaa caa aat cca gac ata gtc atc tat Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr 325 330 335	1008
caa tac atg gat gat ttg tat gta gga tct gac tta gaa ata ggg cag Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln 340 345 350	1056
cat aga aca aaa ata gag gaa ctg aga caa cat ctg ttg agg tgg gga His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly 355 360 365	1104

ttt acc aca cca gac aaa aaa cat cag aaa gaa cct cca ttc ctt tgg	1152
Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp	
370 375 380	
atg ggt tat gaa ctc cat cct gat aaa tgg aca gta cag cct ata gtg	1200
Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val	
385 390 395 400	
ctg cca gaa aag gac agc tgg act gtc aat gac ata cag aaa tta gtg	1248
Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val	
405 410 415	
gga aaa ttg aat tgg gca agt cag att tat gca ggg att aaa gta agg	1296
Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Arg	
420 425 430	
caa tta tgt aaa ctt ctt agg gga acc aaa gca cta aca gaa gta gta	1344
Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Val	
435 440 445	
cca cta aca gaa gaa gca gag cta gaa ctg gca gaa aac agg gag att	1392
Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile	
450 455 460	
cta aaa gaa ccg gta cat gga gtg tat tat gac cca tca aaa gac tta	1440
Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu	
465 470 475 480	
ata gca gaa ata cag aag cag ggg caa ggc caa tgg aca tat caa att	1488
Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile	
485 490 495	
tat caa gag cca ttt aaa aat ctg aaa aca gga aaa tat gca aga atg	1536
Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met	
500 505 510	
aag ggt gcc cac act aat gat gtg aaa caa tta aca gag gca gta caa	1584
Lys Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln	
515 520 525	
aaa ata gcc aca gaa agc ata gta ata tgg gga aag act cct aaa ttt	1632
Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe	
530 535 540	
aaa tta ccc ata caa aag gaa aca tgg gaa gca tgg tgg aca gag tat	1680
Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr Glu Tyr	
545 550 555 560	
tgg caa gcc acc tgg att cct gag tgg gag ttt gtc aat acc cct ccc	1728
Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro	
565 570 575	

tta gtg aag tta tgg tac cag tta gag aaa gaa ccc ata ata gga gca Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ile Gly Ala 580 585 590	1776
gaa act ttc tat gta gat ggg gca gcc aat agg gaa act aaa tta gga Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly 595 600 605	1824
aaa gca gga tat gta act gac aga gga aga caa aaa gtt gtc ccc cta Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu 610 615 620	1872
acg gac aca aca aat cag aag act gag tta caa gca att cat cta gct Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala 625 630 635 640	1920
ttg cag gat tcg gga tta gaa gta aac ata gtg aca gac tca caa tat Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr 645 650 655	1968
gca ttg gga atc att caa gca caa cca gat aag agt gaa tca gag tta Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu 660 665 670	2016
gtc agt caa ata ata gag cag tta ata aaa aag gaa aaa gtc tac ctg Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu 675 680 685	2064
gca tgg gta cca gca cac aaa gga att gga gga aat gaa caa gta gat Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp 690 695 700	2112
ggg ttg gtc agt gct gga atc agg aaa gta cta ttt tta gat gga ata Gly Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile 705 710 715 720	2160
gat aag gcc caa gaa gaa cat gag aaa tat cac agt aat tgg aga gca Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala 725 730 735	2208
atg gct agt gat ttt aac cta cca cct gta gta gca aaa gaa ata gta Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val 740 745 750	2256
gcc agc tgt gat aaa tgt cag cta aaa ggg gaa gcc atg cat gga caa Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln 755 760 765	2304
gta gac tgt agc cca gga ata tgg cag cta gat tgt aca cat tta gaa Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu 770 775 780	2352

gga aaa gtt atc ttg gta gca gtt cat gta gcc agt gga tat ata gaa	2400
Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu	
785 790 795 800	
gca gaa gta att cca gca gag aca ggg caa gaa aca gca tac ttc ctc	2448
Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu	
805 810 815	
tta aaa tta gca gga aga tgg cca gta aaa aca gta cat aca gac aat	2496
Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Val His Thr Asp Asn	
820 825 830	
ggc agc aat ttc acc agt act aca gtt aag gcc gcc tgt tgg tgg gcg	2544
Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala	
835 840 845	
ggg atc aag cag gaa ttt ggc att ccc tac aat ccc caa agt caa gga	2592
Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly	
850 855 860	
gta ata gaa tct atg aat aaa gaa tta aag aaa att ata gga cag gta	2640
Val Ile Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val	
865 870 875 880	
aga gat cag gct gaa cat ctt aag aca gca gta caa atg gca gta ttc	2688
Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe	
885 890 895	
atc cac aat ttt aaa aga aaa ggg ggg att ggg ggg tac agt gca ggg	2736
Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly	
900 905 910	
gaa aga ata gta gac ata ata gca aca gac ata caa act aaa gaa tta	2784
Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu	
915 920 925	
caa aaa caa att aca aaa att caa aat ttt cgg gtt tat tac agg gac	2832
Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp	
930 935 940	
agc aga gat cca gtt tgg aaa gga cca gca aag ctc ctc tgg aaa ggt	2880
Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly	
945 950 955 960	
gaa ggg gca gta gta ata caa gat aat agt gac ata aaa gta gtg cca	2928
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro	
965 970 975	
aga aga aaa gca aag atc atc agg gat tat gga aaa cag atg gca ggt	2976
Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly	
980 985 990	

gat gat tgt gtg gca agt aga cag gat gag gat taa
 Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp *
 995 1000

3012

<210> 7
 <211> 1003
 <212> PRT
 <213> Homo sapien

<400> 7
 Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe
 1 5 10 15
 Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln
 20 25 30
 Val Trp Gly Arg Asp Asn Asn Ser Leu Ser Glu Ala Gly Ala Asp Arg
 35 40 45
 Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg
 50 55 60
 Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
 65 70 75 80
 Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Met Asn Leu Pro Gly
 85 90 95
 Arg Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
 100 105 110
 Gly Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile
 115 120 125
 Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
 130 135 140
 Leu Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
 145 150 155 160
 Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 165 170 175
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
 180 185 190
 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 195 200 205
 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 210 215 220
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 225 230 235 240
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 245 250 255
 Gln Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
 260 265 270
 Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 275 280 285
 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 290 295 300
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr
 305 310 315 320

Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
 325 330 335
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 340 345 350
 His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
 355 360 365
 Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 370 375 380
 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
 385 390 395 400
 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
 405 410 415
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Arg
 420 425 430
 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Val
 435 440 445
 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 450 455 460
 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 465 470 475 480
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
 485 490 495
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
 500 505 510
 Lys Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 515 520 525
 Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr Glu Tyr
 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ile Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu
 610 615 620
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala
 625 630 635 640
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu
 660 665 670
 Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 675 680 685
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700
 Gly Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735

Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780
 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
 805 810 815
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Val His Thr Asp Asn
 820 825 830
 Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala
 835 840 845
 Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 850 855 860
 Val Ile Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
 865 870 875 880
 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 885 890 895
 Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 900 905 910
 Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 915 920 925
 Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 930 935 940
 Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 945 950 955 960
 Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
 965 970 975
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 980 985 990
 Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp
 995 1000

<210> 8
 <211> 3012
 <212> DNA
 <213> Homo sapien

<220>
 <221> CDS
 <222> (1)...(3012)

<400> 8
 ttt ttt agg gaa gat ctg gcc ttc cca caa ggg aag gcc agg gaa ttt 48
 Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe
 1 5 10 15
 tct tca gag cag acc aga gcc aac agc ccc acc aga aga gag ctt cag 96
 Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln
 20 25 30

gtt tgg gga aga gac aac aac tcc ctc tca gaa gca gga gcc gat aga Val Trp Gly Arg Asp Asn Asn Ser Leu Ser Glu Ala Gly Ala Asp Arg 35 40 45	144
caa gga act gta tcc ttt agc ttc cct cag atc act ctt tgg cag cga Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg 50 55 60	192
ccc ctc gtc aca ata aag ata ggg ggg caa tta aag gaa gct cta tta Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu 65 70 75 80	240
gat aca gga gca gat gat aca gta tta gaa gaa atg aat ttg cca gga Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Met Asn Leu Pro Gly 85 90 95	288
aga tgg aaa cca aaa atg ata ggg gga att gga ggt ttt atc aaa gta Arg Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val 100 105 110	336
aga cag tat gat cag ata ctc ata gaa atc tgc gga cat aaa gct ata Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile 115 120 125	384
ggg aca gta tta gta gga cct aca cct gtc aac ata att gga aga aat Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn 130 135 140	432
ctg ttg act cag att ggc tgc act tta aat ttt ccc att agt cct att Leu Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile 145 150 155 160	480
gag act gta cca gta aaa tta aag cca gga atg gat ggc cca aaa gtt Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val 165 170 175	528
aaa caa tgg cca ttg aca gaa gaa aaa ata aaa gca tta gta gaa att Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile 180 185 190	576
tgt aca gaa atg gaa aag gaa gga aaa att tca aaa att ggg cct gaa Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu 195 200 205	624
aat cca tac aat act cca gta ttt gcc ata aag aaa aaa gac agt act Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr 210 215 220	672
aaa tgg aga aaa tta gta gat ttc aga gaa ctt aat aag aga act caa Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln 225 230 235 240	720

gat ttc tgg gaa gtt caa tta gga ata cca cat cct gca ggg tta aaa Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys 245 250 255	768
cag aaa aaa tca gta aca gta ctg gat gtg ggc gat gca tat ttt tca Gln Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser 260 265 270	816
gtt ccc tta gat aaa gac ttc agg aag tat act gca ttt acc ata cct Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro 275 280 285	864
agt ata aac aat gag aca cca ggg att aga tat cag tac aat gtg ctt Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu 290 295 300	912
cca cag gga tgg aaa gga tca cca gca ata ttc cag tgt agc atg aca Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr 305 310 315 320	960
aaa atc tta gag cct ttt aga aaa caa aat cca gac ata gtc atc tat Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr 325 330 335	1008
caa tac atg gat gat ttg tat gta gga tct gac tta gaa ata ggg cag Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln 340 345 350	1056
cat aga aca aaa ata gag gaa ctg aga caa cat ctg ttg agg tgg gga His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly 355 360 365	1104
ttt acc aca cca gac aaa aaa cat cag aaa gaa cct cca ttc ctt tgg Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp 370 375 380	1152
atg ggt tat gaa ctc cat cct gat aaa tgg aca gta cag cct ata gtg Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val 385 390 395 400	1200
ctg cca gaa aag gac agc tgg act gtc aat gac ata cag aaa tta gtg Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val 405 410 415	1248
gga aaa ttg aat tgg gca agt cag att tat gca ggg att aaa gta agg Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Arg 420 425 430	1296
caa tta tgt aaa ctt ctt agg gga acc aaa gca cta aca gaa gta gta Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Val 435 440 445	1344

cca cta aca gaa gaa gca gag cta gaa ctg gca gaa aac agg gag att	1392
Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile	
450 455 460	
cta aaa gaa ccg gta cat gga gtg tat tat gac cca tca aaa gac tta	1440
Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu	
465 470 475 480	
ata gca gaa ata cag aag cag ggg caa ggc caa tgg aca tat caa att	1488
Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile	
485 490 495	
tat caa gag cca ttt aaa aat ctg aaa aca gga aaa tat gca aga atg	1536
Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met	
500 505 510	
aag ggt gcc cac act aat gat gtg aaa caa tta aca gag gca gta caa	1584
Lys Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln	
515 520 525	
aaa ata gcc aca gaa agc ata gta ata tgg gga aag act cct aaa ttt	1632
Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe	
530 535 540	
aaa tta ccc ata caa aag gaa aca tgg gaa gca tgg tgg aca gag tat	1680
Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr Glu Tyr	
545 550 555 560	
tgg caa gcc acc tgg att cct gag tgg gag ttt gtc aat acc cct ccc	1728
Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro	
565 570 575	
tta gtg aag tta tgg tac cag tta gag aaa gaa ccc ata ata gga gca	1776
Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ile Gly Ala	
580 585 590	
gaa act ttc tat gta gat ggg gca gcc aat agg gaa act aaa tta gga	1824
Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly	
595 600 605	
aaa gca gga tat gta act gac aga gga aga caa aaa gtt gtc ccc cta	1872
Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu	
610 615 620	
acg gac aca aca aat cag aag act gag tta caa gca att cat cta gct	1920
Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala	
625 630 635 640	
ttg cag gat tcg gga tta gaa gta aac ata gtg aca gac tca caa tat	1968
Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr	
645 650 655	

gca ttg gga atc att caa gca caa cca gat aag agt gaa tca gag tta Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu 660 665 670	2016
gtc agt caa ata ata gag cag tta ata aaa aag gaa aaa gtc tac ctg Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu 675 680 685	2064
gca tgg gta cca gca cac aaa gga att gga gga aat gaa caa gta gat Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp 690 695 700	2112
aag ttg gtc agt gct gga atc agg aaa gta cta ttt tta gat gga ata Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile 705 710 715 720	2160
gat aag gcc caa gaa gaa cat gag aaa tat cac agt aat tgg aga gca Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala 725 730 735	2208
atg gct agt gat ttt aac cta cca cct gta gta gca aaa gaa ata gta Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val 740 745 750	2256
gcc agc tgt gat aaa tgt cag cta aaa ggg gaa gcc atg cat gga caa Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln 755 760 765	2304
gta gac tgt agc cca gga ata tgg cag cta gat tgt aca cat tta gaa Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu 770 775 780	2352
gga aaa gtt atc ttg gta gca gtt cat gta gcc agt gga tat ata gaa Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu 785 790 795 800	2400
gca gaa gta att cca gca gag aca ggg caa gaa aca gca tac ttc ctc Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu 805 810 815	2448
tta aaa tta gca gga aga tgg cca gta aaa aca gta cat aca gac aat Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Val His Thr Asp Asn 820 825 830	2496
ggc agc aat ttc acc agt act aca gtt aag gcc gcc tgt tgg tgg gcg Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala 835 840 845	2544
ggg atc aag cag gaa ttt ggc att ccc tac aat ccc caa agt caa gga Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly 850 855 860	2592

gta ata gaa tct atg aat aaa gaa tta aag aaa att ata gga cag gta 2640
Val Ile Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
865 870 875 880

aga gat cag gct gaa cat ctt aag aca gca gta caa atg gca gta ttc 2688
Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
885 890 895

atc cac aat ttt aaa aga aaa ggg ggg att ggg ggg tac agt gca ggg 2736
Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
900 905 910

gaa aga ata gta gac ata ata gca aca gac ata caa act aaa gaa tta 2784
Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
915 920 925

caa aaa caa att aca aaa att caa aat ttt cgg gtt tat tac agg gac 2832
Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
930 935 940

agc aga gat cca gtt tgg aaa gga cca gca aag ctc ctc tgg aaa ggt 2880
Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
945 950 955 960

gaa ggg gca gta gta ata caa gat aat agt gac ata aaa gta gtg cca 2928
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
965 970 975

aga aga aaa gca aag atc atc agg gat tat gga aaa cag atg gca ggt 2976
Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
980 985 990

gat gat tgt gtg gca agt aga cag gat gag gat taa 3012
Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp *
995 1000

<210> 9

<211> 1003

<212> PRT

<213> Homo sapien

<400> 9

Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe
1 5 10 15
Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln
20 25 30
Val Trp Gly Arg Asp Asn Asn Ser Leu Ser Glu Ala Gly Ala Asp Arg
35 40 45
Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg
50 55 60

Pro	Leu	Val	Thr	Ile	Lys	Ile	Gly	Gly	Gln	Leu	Lys	Glu	Ala	Leu	Leu	65	70	75	80
Asp	Thr	Gly	Ala	Asp	Asp	Thr	Val	Leu	Glu	Glu	Met	Asn	Leu	Pro	Gly	85	90	95	
Arg	Trp	Lys	Pro	Lys	Met	Ile	Gly	Gly	Ile	Gly	Gly	Phe	Ile	Lys	Val	100	105	110	
Arg	Gln	Tyr	Asp	Gln	Ile	Leu	Ile	Glu	Ile	Cys	Gly	His	Lys	Ala	Ile	115	120	125	
Gly	Thr	Val	Leu	Val	Gly	Pro	Thr	Pro	Val	Asn	Ile	Ile	Gly	Arg	Asn	130	135	140	
Leu	Leu	Thr	Gln	Ile	Gly	Cys	Thr	Leu	Asn	Phe	Pro	Ile	Ser	Pro	Ile	145	150	155	160
Glu	Thr	Val	Pro	Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro	Lys	Val	165	170	175	
Lys	Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val	Glu	Ile	180	185	190	
Cys	Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys	Ile	Ser	Lys	Ile	Gly	Pro	Glu	195	200	205	
Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	Ser	Thr	210	215	220	
Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln	225	230	235	240
Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys	245	250	255	
Gln	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser	260	265	270	
Val	Pro	Leu	Asp	Lys	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro	275	280	285	
Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu	290	295	300	
Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Cys	Ser	Met	Thr	305	310	315	320
Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	Ile	Tyr	325	330	335	
Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln	340	345	350	
His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	Trp	Gly	355	360	365	
Phe	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp	370	375	380	
Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Val	385	390	395	400
Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val	405	410	415	
Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Ala	Gly	Ile	Lys	Val	Arg	420	425	430	
Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	Val	Val	435	440	445	
Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile	450	455	460	
Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu	465	470	475	480

Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	Gln	Ile	485	490	495
Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Arg	Met	500	505	510
Lys	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	Val	Gln	515	520	525
Lys	Ile	Ala	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	Lys	Phe	530	535	540
Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Ala	Trp	Trp	Thr	Glu	Tyr	545	550	555
Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	Pro	Pro	565	570	575
Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu	Pro	Ile	Ile	Gly	Ala	580	585	590
Glu	Thr	Phe	Tyr	Val	Asp	Gly	Ala	Ala	Asn	Arg	Glu	Thr	Lys	Leu	Gly	595	600	605
Lys	Ala	Gly	Tyr	Val	Thr	Asp	Arg	Gly	Arg	Gln	Lys	Val	Val	Pro	Leu	610	615	620
Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Glu	Leu	Gln	Ala	Ile	His	Leu	Ala	625	630	635
Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Asp	Ser	Gln	Tyr	645	650	655
Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Lys	Ser	Glu	Ser	Glu	Leu	660	665	670
Val	Ser	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys	Val	Tyr	Leu	675	680	685
Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln	Val	Asp	690	695	700
Lys	Leu	Val	Ser	Ala	Gly	Ile	Arg	Lys	Val	Leu	Phe	Leu	Asp	Gly	Ile	705	710	715
Asp	Lys	Ala	Gln	Glu	Glu	His	Glu	Lys	Tyr	His	Ser	Asn	Trp	Arg	Ala	725	730	735
Met	Ala	Ser	Asp	Phe	Asn	Leu	Pro	Pro	Val	Val	Ala	Lys	Glu	Ile	Val	740	745	750
Ala	Ser	Cys	Asp	Lys	Cys	Gln	Leu	Lys	Gly	Glu	Ala	Met	His	Gly	Gln	755	760	765
Val	Asp	Cys	Ser	Pro	Gly	Ile	Trp	Gln	Leu	Asp	Cys	Thr	His	Leu	Glu	770	775	780
Gly	Lys	Val	Ile	Leu	Val	Ala	Val	His	Val	Ala	Ser	Gly	Tyr	Ile	Glu	785	790	795
Ala	Glu	Val	Ile	Pro	Ala	Glu	Thr	Gly	Gln	Glu	Thr	Ala	Tyr	Phe	Leu	805	810	815
Leu	Lys	Leu	Ala	Gly	Arg	Trp	Pro	Val	Lys	Thr	Val	His	Thr	Asp	Asn	820	825	830
Gly	Ser	Asn	Phe	Thr	Ser	Thr	Thr	Val	Lys	Ala	Ala	Cys	Trp	Trp	Ala	835	840	845
Gly	Ile	Lys	Gln	Glu	Phe	Gly	Ile	Pro	Tyr	Asn	Pro	Gln	Ser	Gln	Gly	850	855	860
Val	Ile	Glu	Ser	Met	Asn	Lys	Glu	Leu	Lys	Lys	Ile	Ile	Gly	Gln	Val	865	870	875
Arg	Asp	Gln	Ala	Glu	His	Leu	Lys	Thr	Ala	Val	Gln	Met	Ala	Val	Phe	885	890	895


```

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
      900                      905                      910
Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
      915                      920                      925
Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
      930                      935                      940
Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
945                      950                      955                      960
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
      965                      970                      975
Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
      980                      985                      990
Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp
      995                      1000

```

```

<210> 10
<211> 3012
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Codon optimized form of HIV pol coding region

<221> CDS
<222> (1) ... (3012)

```

```

<400> 10
ttt ttt agg gaa gat ctg gcc ttc cca caa ggg aag gcc agg gaa ttt      48
Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe
  1                      5                      10                      15

tct tca gag cag acc aga gcc aac agc ccc acc aga aga gag ctt cag      96
Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln
      20                      25                      30

gtt tgg gga aga gac aac aac tcc ctc tca gaa gca gga gcc gat aga      144
Val Trp Gly Arg Asp Asn Asn Ser Leu Ser Glu Ala Gly Ala Asp Arg
      35                      40                      45

caa gga act gta tcc ttt agc ttc cct cag atc act ctt tgg cag cga      192
Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg
      50                      55                      60

ccc ctc gtc aca ata aag atc ggt ggc cag ctg aag gag gcc ctg ctg      240
Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
      65                      70                      75                      80

gac acc ggc gcc gac gac acc gtg ctg gag gag atg aac ctg ccc ggc      288
Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Met Asn Leu Pro Gly
      85                      90                      95

```

cgc tgg aag ccc aag atg atc ggc ggc atc ggc ggc ttc atc aaa gtc	336
Arg Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val	
100 105 110	
cgc cag tac gac cag atc ctg atc gag atc tgc ggc cac aag gcc atc	384
Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile	
115 120 125	
ggc acc gtg ctg gtg ggc ccc acc ccc gtg aac atc atc ggc cgc aac	432
Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn	
130 135 140	
ctg ctg acc cag atc ggc tgc acc ctg aac ttc ccc atc tcc ccc atc	480
Leu Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile	
145 150 155 160	
gag acc gtg ccc gtg aag ctg aag ccc ggc atg gac ggc ccc aaa gtc	528
Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val	
165 170 175	
aag cag tgg ccc ctg acc gag gag aag atc aag gcc ctg gtg gag atc	576
Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile	
180 185 190	
tgc acc gag atg gag aag gag ggc aag atc tcc aag atc ggc ccc gag	624
Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu	
195 200 205	
aac ccc tac aac acc ccc gtg ttc gcc atc aag aag aag gac tcc acc	672
Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr	
210 215 220	
aag tgg cgc aag ctg gtg gac ttc cgc gag ctg aac aag cgc acc cag	720
Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln	
225 230 235 240	
gac ttc tgg gag gtg cag ctg ggc atc ccc cac ccc gcc ggc ctg aag	768
Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys	
245 250 255	
cag aag aag tcc gtg acc gtg ctg gac gtg ggc gac gcc tac ttc tcc	816
Gln Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser	
260 265 270	
gtg ccc ctg gac aag gac ttc cgc aag tac acc gcc ttc acc atc ccc	864
Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro	
275 280 285	
tcc atc aac aac gag acc ccc ggc atc cgc tac cag tac aac gtg ctg	912
Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu	
290 295 300	

ccc cag ggc tgg aag ggc tcc ccc gcc atc ttc cag tgc tcc atg acc	960
Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr	
305 310 315 320	
aag atc ctg gag ccc ttc cgc aag cag aac ccc gac atc gtg atc tac	1008
Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr	
325 330 335	
cag tac atg gac gac ctg tac gtg ggc tcc gac ctg gag atc ggc cag	1056
Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln	
340 345 350	
cac cgc acc aag atc gag gag ctg cgc cag cac ctg ctg cgc tgg ggc	1104
His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly	
355 360 365	
ttc acc acc ccc gac aag aag cac cag aag gag ccc ccc ttc ctg tgg	1152
Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp	
370 375 380	
atg ggc tac gag ctg cac ccc gac aag tgg acc gtg cag ccc atc gtg	1200
Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val	
385 390 395 400	
ctg ccc gag aag gac tcc tgg acc gtg aac gac atc cag aag ctg gtg	1248
Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val	
405 410 415	
ggc aag ctg aac tgg gcc tcc cag atc tac gcc ggc atc aaa gtc cgc	1296
Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Arg	
420 425 430	
cag ctg tgc aag ctg ctg cgc ggc acc aag gcc ctg acc gag gtg gtg	1344
Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Val	
435 440 445	
ccc ctg acc gag gag gcc gag ctg gag ctg gcc gag aac cgc gag atc	1392
Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile	
450 455 460	
ctg aag gag ccc gtg cac ggc gtg tac tac gac ccc tcc aag gac ctg	1440
Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu	
465 470 475 480	
atc gcc gag atc cag aag cag ggc cag ggc cag tgg acc tac cag atc	1488
Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile	
485 490 495	
tac cag gag ccc ttc aag aac ctg aag acc ggc aaa tac gcc cgc atg	1536
Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met	
500 505 510	

aag ggc gcc cac acc aac gac gtg aag cag ctg acc gag gcc gtg cag Lys Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln 515 520 525	1584
aag atc gcc acc gag tcc atc gtg atc tgg ggc aag act ccc aag ttc Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe 530 535 540	1632
aag ctg ccc atc cag aag gag acc tgg gag gcc tgg tgg acc gag tac Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr Glu Tyr 545 550 555 560	1680
tgg cag gcc acc tgg atc ccc gag tgg gag ttc gtg aac acc ccc ccc Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro 565 570 575	1728
ctg gtg aag ctg tgg tac cag ctg gag aag gag ccc atc atc ggc gcc Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ile Gly Ala 580 585 590	1776
gag acc ttc tac gtg gac ggc gcc gcc aac cgc gag acc aag ctg ggc Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly 595 600 605	1824
aag gcc ggc tac gtg acc gac cgc ggc cgc cag aag gtg gtg ccc ctg Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu 610 615 620	1872
acc gac acc acc aac cag aag acc gag ctg cag gcc atc cac ctg gcc Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala 625 630 635 640	1920
ctg caa gac tcc ggc ctg gag gtg aac atc gtg acc gac tcc cag tat Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr 645 650 655	1968
gca ttg ggc atc atc cag gcc cag ccc gac aag tcc gag tcc gag ctg Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu 660 665 670	2016
gtg tcc cag atc atc gag cag ctg atc aag aag gag aag gtg tac ctg Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu 675 680 685	2064
gcc tgg gtg ccc gcc cac aag ggc atc ggc ggc aac gag cag gtg gac Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp 690 695 700	2112
aag ctg gtg tcc gcc ggc atc cgc aag gtg ctg ttc ctg gac ggc atc Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile 705 710 715 720	2160

gac aag gcc cag gag gag cac gag aag tac cac tcc aac tgg cgc gcc Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala 725 730 735	2208
atg gcc tcc gac ttc aac ctg ccc ccc gtg gtg gcc aag gag atc gtg Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val 740 745 750	2256
gcc tcc tgc gac aag tgc cag ctg aag ggc gag gcc atg cac ggc cag Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln 755 760 765	2304
gtg gac tgc tcc ccc ggc atc tgg cag ctg gac tgc acc cac ctg gag Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu 770 775 780	2352
ggc aag gtg atc ctg gtg gcc gtg cac gtg gcc tcc ggc tac atc gag Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu 785 790 795 800	2400
gcc gag gtg atc ccc gcc gag acc ggc cag gag acc gcc tac ttc ctg Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu 805 810 815	2448
ctg aag ctg gcc ggc cgc tgg ccc gtg aag acc gtg cac acc gac aac Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Val His Thr Asp Asn 820 825 830	2496
ggc tcc aac ttc acc tcc acc acc gtg aag gcc gcc tgc tgg tgg gcc Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala 835 840 845	2544
ggc atc aag cag gag ttc ggc atc ccc tac aac ccc cag tcc cag ggc Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly 850 855 860	2592
gtg atc gag tcc atg aac aag gag ctg aag aag atc atc ggc caa gtc Val Ile Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val 865 870 875 880	2640
cgc gac cag gcc gag cac ctg aag acc gcc gtg cag atg gcc gtg ttc Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe 885 890 895	2688
atc cac aac ttc aag cgc aag ggc ggc atc ggc ggc tac tcc gcc ggc Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly 900 905 910	2736
gag cgc atc gtg gac atc atc gcc acc gac atc cag acc aag gag ctg Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu 915 920 925	2784

```

cag aag cag atc acc aag atc cag aac ttc cgc gtg tac tac cgc gac      2832
Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
   930                      935                      940

tcc cgc gac ccc gtg tgg aag ggc ccc gcc aag ctg ctg tgg aag ggc      2880
Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
   945                      950                      955                      960

gag ggc gcc gtg gtg atc cag gac aac tcc gac atc aag gtg gtg ccc      2928
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
                   965                      970                      975

cgc cgc aag gcc aag atc atc cgc gac tac ggc aag cag atg gcc ggc      2976
Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Lys Gln Met Ala Gly
                   980                      985                      990

gac gac tgc gtg gcc tcc cgc cag gac gag gac taa      3012
Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp *
   995                      1000

```

```

<210> 11
<211> 1003
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Codon optimized form of HIV pol coding region

```

```

<400> 11
Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe
 1             5             10             15
Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln
                20             25             30
Val Trp Gly Arg Asp Asn Asn Ser Leu Ser Glu Ala Gly Ala Asp Arg
   35             40             45
Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg
   50             55             60
Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
 65             70             75             80
Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Met Asn Leu Pro Gly
                85             90             95
Arg Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
                100            105            110
Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile
                115            120            125
Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
 130            135            140
Leu Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
 145            150            155            160
Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
                165            170            175

```

Lys	Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val	Glu	Ile
			180					185					190		
Cys	Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys	Ile	Ser	Lys	Ile	Gly	Pro	Glu
		195					200					205			
Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	Ser	Thr
	210					215					220				
Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln
225					230					235					240
Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys
			245					250						255	
Gln	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser
			260					265					270		
Val	Pro	Leu	Asp	Lys	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro
		275					280					285			
Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu
	290					295					300				
Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Cys	Ser	Met	Thr
305					310					315					320
Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	Ile	Tyr
			325						330					335	
Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln
			340					345					350		
His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	Trp	Gly
		355					360					365			
Phe	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp
		370				375					380				
Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Val
385					390					395					400
Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val
			405						410					415	
Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Ala	Gly	Ile	Lys	Val	Arg
			420					425					430		
Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	Val	Val
		435					440					445			
Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile
		450				455					460				
Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu
465					470					475					480
Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	Gln	Ile
			485						490					495	
Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Arg	Met
			500					505					510		
Lys	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	Val	Gln
		515					520					525			
Lys	Ile	Ala	Thr	Glu											

Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu
 610 615 620
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala
 625 630 635 640
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu
 660 665 670
 Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 675 680 685
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700
 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780
 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
 805 810 815
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Val His Thr Asp Asn
 820 825 830
 Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala
 835 840 845
 Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 850 855 860
 Val Ile Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
 865 870 875 880
 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 885 890 895
 Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 900 905 910
 Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 915 920 925
 Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 930 935 940
 Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 945 950 955 960
 Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
 965 970 975
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 980 985 990
 Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp
 995 1000

<210> 12
 <211> 8908
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Packaging construct pHDmHgpm2 comprising a codon
 optimized form of HIV gag pol region

<400> 12

agcttgcccc	attgcatacg	ttgtatccat	atcataatat	gtacatttat	attggctcat	60
gtccaacatt	accgccatgt	tgacattgat	tattgactag	ttattaatag	taatcaatta	120
cggggctcatt	agttcatagc	ccatatatgg	agttccgcgt	tacataactt	acggtaaatg	180
gcccgccttg	ctgaccgccc	aacgaccccc	gccattgac	gtcaataatg	acgtatgttc	240
ccatagtaac	gccaataggg	actttccatt	gacgtcaatg	ggtggagtat	ttacggtaaa	300
ctgcccactt	ggcagtacat	caagtgtatc	atatgccaa	tacgccccct	attgacgtca	360
atgacggtaa	atggcccgc	tggcattatg	cccagtacat	gaccttatgg	gactttccta	420
cttggcagta	catctacgta	ttagtcacg	ctattaccat	ggtgatgcgg	ttttggcagt	480
acatcaatgg	gcgtggatag	cggtttgact	cacggggatt	tccaagtctc	caccccattg	540
acgtcaatgg	gagtttgatt	tggcaccaaa	atcaacggga	ctttccaaaa	tgtcgtaaca	600
actccgcccc	attgacgcaa	atgggcggta	ggcgtgtacg	gtgggaggtc	tatataagca	660
gagctcgttt	agtgaaccgt	cagatcgcc	ggagacgcca	tccacgctgt	tttgacctcc	720
atagaagaca	ccgggaccga	tccagcctcc	cctcgaagct	gatcctgaga	acttcagggt	780
gagtctatgg	gacccttgat	gttttctttc	cccttctttt	ctatggttaa	gttcatgtca	840
taggaagggg	agaagtaaca	gggtacacat	attgaccaaa	tcagggtaat	tttgcatattg	900
taatttttaa	aaatgctttc	ttcttttaat	atactttttt	gtttatctta	tttctaatac	960
tttcctaat	ctctttcttt	cagggaata	atgatacaat	gtatcatgcc	tctttgcacc	1020
attctaaaga	ataacagtga	taatttctgg	gttaaggcaa	tagcaatatt	tctgcatata	1080
aatatttctg	catataaatt	gtaactgatg	taagagggtt	catattgcta	atagcagcta	1140
caatccagct	accattctgc	ttttatttta	tgggtgggat	aaggctggat	tattctgagt	1200
ccaagctagg	cccttttgct	aatcatgttc	atacctctta	tcttctctcc	acagctcctg	1260
ggcaacgtgc	tggctctgtg	gctggcccat	cactttggca	agaattctta	gactgccatg	1320
ggcgcccgcg	cctccgtgct	gtccggcggc	gagctggaca	agtgggagaa	gatccgcctg	1380
cgccccggcg	gcaagaagca	gtacaagctg	aagcacatcg	tgtgggcctc	ccgcgagctg	1440
gagcgcttcg	ccgtgaaccc	cggcctgctg	gagacctccg	agggctgccg	ccagatcctg	1500
ggccagctgc	agccctccct	gcaaaccggc	tccgaggagc	tgcgtccct	gtacaacacc	1560
atcgccgtgc	tgtactgctg	gcaccagcgc	atcgacgtga	aggacaccaa	ggaggccctg	1620
gacaagatcg	aggaggagca	gaacaagtcc	aagaagaagg	cccagcaggc	cgccgccgac	1680
accggcaaca	actcccaggt	gtcccagaac	taccccatcg	tgcagaacct	gcagggccag	1740
atggtgcacc	aggccatctc	cccccgacc	ctgaacgcct	gggtgaaggt	ggtggaggag	1800
aaggccttct	ccccgaagt	catcccatg	ttctccgccc	tgtccgaggg	cgccaccccc	1860
caggacctga	acaccatgct	gaacaccgtg	ggcgccacc	aggccgccat	gcagatgctg	1920
aaggagacca	tcaacgagga	ggccgccgag	tgggaccgcc	tgcaccccg	gcacgccggc	1980
cccatcgccc	ccggccagat	gcgcgagccc	cgcggctccg	acatcgccgg	caccacctcc	2040
accctgcaag	agcagatcgg	ctggatgacc	cacaaccccc	ccatccccgt	gggcgagatc	2100
tacaagcgct	ggatcatcct	gggcctgaac	aagatcgtgc	gcatgtactc	ccccacctcc	2160
atcctggaca	tccgccaggg	ccccaaaggag	cccttccg	actacgtgga	ccgcttctac	2220
aagacctg	gcgcgagca	ggcctccag	gaggtaaaga	actggatgac	cgagacctg	2280
ctggtgcaga	acgccaaccc	cgactgcaag	accatcctga	aggccctggg	ccccggcgcc	2340
accctggagg	agatgatgac	cgcctgccag	ggcgtggg	gccccggcca	caaggcccgc	2400
gtgctggccg	aggccatgtc	ccaagtcacc	aaccccgcca	ccatcatgat	ccagaagggc	2460
aacttccgca	accagcgcaa	gaccgtgaag	tgttcaact	gcggcaagga	gggccacatc	2520

gccaagaact	gccgcgcccc	ccgcaagaag	ggctgctgga	agtgcggcaa	ggagggccac	2580
cagatgaaag	attgtactga	gagacaggct	aatttttttag	ggaagatctg	gccttccccc	2640
aaggggaaggc	caggggaattt	tcttcagagc	agaccagagc	caacagcccc	accagaagag	2700
agcttcaggt	ttggggaaga	gacaacaact	ccctctcaga	agcaggagcc	gatagacaag	2760
gaactgtatc	ctttagcttc	cctcagatca	ctctttggca	gcgacccctc	gtcacaataa	2820
agatcggtgg	ccagctgaag	gaggccctgc	tggacaccgg	cgccgacgac	accgtgctgg	2880
aggagatgaa	cctgcccggc	cgctggaagc	ccaagatgat	cggcggcatc	ggcggcttca	2940
tcaaagtccg	ccagtacgac	cagatcctga	tcgagatctg	cggccacaag	gccatcggca	3000
ccgtgctggt	gggccccacc	cccgtgaaca	tcacgggccg	caacctgctg	accagatcg	3060
gctgcaccct	gaacttcccc	atctccccca	tcgagaccgt	gcccgtgaag	ctgaagcccc	3120
gcatggacgg	ccccaaagtc	aagcagtggc	ccctgaccga	ggagaagatc	aaggccctgg	3180
tggagatctg	caccgagatg	gagaaggagg	gcaagatctc	caagatcggc	cccagagaacc	3240
cctacaacac	ccccgtgttc	gccatcaaga	agaaggactc	caccaagtgg	cgcaagctgg	3300
tggacttccg	cgagctgaac	aagcgcaccc	aggacttctg	ggaggtgcag	ctgggcatcc	3360
cccacccccg	cggcctgaag	cagaagaagt	ccgtgaccgt	gctggacgtg	ggcgacgcct	3420
acttctccgt	gcccctggac	aaggacttcc	gcaagtacac	cgccttcacc	atccccctca	3480
tcaacaacga	gacccccggc	atccgctacc	agtacaacgt	gctgccccag	ggctggaagg	3540
gctcccccg	catcttccag	tgctccatga	ccaagatcct	ggagcccttc	cgcaagcaga	3600
accccgacat	cgtgatctac	cagtacatgg	acgacctgta	cgtgggctcc	gacctggaga	3660
tcggccagca	ccgcaccaag	atcgaggagc	tgcgccagca	cctgctgcgc	tggggcttca	3720
ccacccccga	caagaagcac	cagaaggagc	cccccttcct	gtggatgggc	tacgagctgc	3780
accccgacaa	gtggaccgtg	cagcccatcg	tgctgcccga	gaaggactcc	tggaccgtga	3840
acgacatcca	gaagctgggt	ggcaagctga	actgggcctc	ccagatctac	gccggcatca	3900
aagtccgcca	gctgtgcaag	ctgctgcgcg	gcaccaaggc	cctgaccgag	gtggtgcccc	3960
tgaccgagga	ggccgagctg	gagctggccg	agaaccgcca	gatcctgaag	gagcccgtgc	4020
acggcggtga	ctacgacccc	tccaaggacc	tgatcgccga	gatccagaag	cagggccagg	4080
gccagtggac	ctaccagatc	taccaggagc	ccttcaagaa	cctgaagacc	ggcaaatacg	4140
ccgcgatgaa	gggcgcccac	accaacgacg	tgaagcagct	gaccgaggcc	gtgcagaaga	4200
tcgccaccga	gtccatcgtg	atctggggca	agactcccaa	gttcaagctg	cccatccaga	4260
aggagacctg	ggaggcctgg	tggaccgagt	actggcaggc	cacctggatc	cccaggtggg	4320
agttcgtgaa	cacccccccc	ctggtgaagc	tgtggtacca	gctggagaag	gagcccatca	4380
tcggcgccga	gaccttctac	gtggacggcg	ccgccaaccg	cgagaccaag	ctgggcaagg	4440
ccggctacgt	gaccgaccgc	ggccgccaga	aggtggtgcc	cctgaccgac	accaccaacc	4500
agaagaccga	gctgcaggcc	atccacctgg	ccctgcaaga	ctccggcctg	gaggtgaaca	4560
tcgtgaccga	ctcccagtat	gcattgggca	tcattccaggc	ccagcccgac	aagtccgagt	4620
ccgagctggt	gtcccagatc	atcgagcagc	tgatcaagaa	ggagaagggtg	tacctggcct	4680
gggtgccccg	ccacaagggc	atcggcggca	acgagcaggt	ggacaagctg	gtgtccgcgc	4740
gcatccgcaa	ggtgctgttc	ctggacggca	tcgacaaggc	ccaggaggag	cacgagaagt	4800
accactccaa	ctggcgcgcc	atggcctccg	acttcaacct	gccccccgtg	gtggccaagg	4860
agatcgtggc	ctcctgcgac	aagtgccagc	tgaaggggcg	ggccatgcac	ggccagggtg	4920
actgctcccc	cggcatctgg	cagctggact	gcacccacct	ggagggcaag	gtgatcctgg	4980
tggcctgca	cgtggcctcc	ggctacatcg	aggccgaggt	gatccccgcc	gagaccggcc	5040
aggagaccgc	ctacttctctg	ctgaagctgg	ccggccgctg	gcccgtgaag	accgtgcaca	5100
ccgacaacgg	ctccaacttc	acctccacca	ccgtgaaggc	cgctgctgg	tgggcccggca	5160
tcaagcagga	gttcggcatc	ccctacaacc	cccagtccca	gggcgtgatc	gagtcctatga	5220
acaaggagct	gaagaagatc	atcggcgaag	tccgcgacca	ggccgagcac	ctgaagaccg	5280
ccgtgcagat	ggccgtgttc	atccacaact	tcaagcgcaa	gggcggcatc	ggcggctact	5340
ccgccggcga	gcgcacgtg	gacatcatcg	ccaccgacat	ccagaccaag	gagctgcaga	5400
agcagatcac	caagatccag	aacttccgcg	tgtactaccg	cgactcccgc	gaccccggtg	5460
ggaagggccc	cgccaagctg	ctgtggaagg	gcgagggcgc	cgtggtgatc	caggacaact	5520
ccgacatcaa	ggtggtgccc	cgccgcaagg	ccaagatcat	ccgcgactac	ggcaagcaga	5580
tggccggcga	cgactgcgtg	gcctcccgc	aggacgagga	ctaacacatg	gaaaagatta	5640

gtaaaacacc	ataggccgct	ctagaggatc	caagcttatc	gataccgctc	acctcgaggg	5700
cccagatcta	attcacccca	ccagtgcagg	ctgcctatca	gaaagtgggtg	gctgggtgtgg	5760
ctaattgccct	ggcccacaag	tatcactaag	ctcgctttct	tgctgtccaa	tttctattaa	5820
aggttccctt	gttccctaag	tccaactact	aaactggggg	atattatgaa	gggccttgag	5880
catctggatt	ctgcctaata	aaaaacattt	attttcattg	caatgatgta	tttaaattat	5940
ttctgaatat	tttactaaaa	agggaaatgtg	ggagggtcagt	gcatttataaa	cataaagaaa	6000
tgaagagcta	gttcaaacct	tgggaaaata	cactatatct	taaactccat	gaaagaaggt	6060
gaggctgcaa	acagctaattg	cacattggca	acagcccctg	atgcctatgc	cttattcatc	6120
cctcagaaaa	ggattcaagt	agaggcttga	tttggagggt	aaagttttgc	tatgctgtat	6180
tttacattac	ttattgtttt	agctgtcctc	atgaatgtct	tttactacc	catttgctta	6240
tcttgcctct	ctcagccttg	actccactca	gttctcttgc	ttagagatac	cacctttccc	6300
ctgaagtgtt	ccttccatgt	tttacggcga	gatgggttct	cctcgccctgg	ccactcagcc	6360
ttagtgtgtc	ctgttgtctt	atagagggtct	acttgaagaa	ggaaaaacag	ggggcatggt	6420
ttgactgtcc	tgtgagccct	tcttccctgc	ctccccact	cacagtgacc	cggaatccct	6480
cgacatggca	gtctagatca	ttcttgaaga	cgaaagggcc	tcgtgatacg	cctattttta	6540
taggttaatg	tcatgataat	aatgggttct	tagacgtcag	gtggcacttt	tcggggaaat	6600
gtgcgcggaa	cccctatttg	tttatttttc	taaatacatt	caaatatgta	tccgctcatg	6660
agacaataac	cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	6720
catttccgtg	tcgcccttat	tccctttttt	gcggcatttt	gccttctctg	ttttgtcac	6780
ccagaaacgc	tgggtgaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	6840
atcgaactgg	atctcaacag	cggtaagatc	cttgagagtt	ttcgccccga	agaacgtttt	6900
ccaatgatga	gcacttttaa	agttctgcta	tgtggcgcg	tattatccc	tattgacgcc	6960
gggcaagagc	aactcggctc	ccgcatacac	tattctcaga	atgacttgg	tgagtactca	7020
ccagtccacg	aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	cagtgtctgc	7080
ataaccatga	gtgataaac	tgcggccaac	ttacttctga	caacgatcgg	aggaccgaag	7140
gagctaaccg	cttttttgca	caacatgggg	gatcatgtaa	ctcgccctga	tcgttgggaa	7200
ccggagctga	atgaagccat	accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	7260
gcaacaacgt	tcgcgaaaact	attaactggc	gaactactta	ctctagcttc	ccggcaacaa	7320
ttaatagact	ggatggaggc	ggataaaagt	gcaggaccac	ttctgcgctc	ggcccttccg	7380
gctggctgg	ttattgtctga	taaatctgga	gccggtgagc	gtgggtctcg	cggtatcatt	7440
gcagcactgg	ggccagatgg	taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	7500
caggcaacta	tggatgaacg	aaatagacag	atcgctgaga	taggtgcctc	actgattaag	7560
cattggtaac	tgtcagacca	agtttactca	tatatacttt	agattgat	aaaacttcat	7620
ttttaattta	aaaggatcta	ggtgaagatc	cttttttgata	atctcatgac	caaaatccct	7680
taacgtgagt	tttcgttcca	ctgagcgtca	gaccccgtag	aaaagatcaa	aggatcttct	7740
tgagatcctt	tttttctg	cgtaatctgc	tgcttgcaaa	caaaaaaacc	accgctacca	7800
gcggtgggtt	gtttgccgga	tcaagagcta	ccaactcttt	ttccgaagg	aactggcttc	7860
agcagagcgc	agataccaaa	tactgttctt	ctagtgtagc	cgtagttagg	ccaccacttc	7920
aagaactctg	tagcaccgcc	tacatacctc	gctctgctaa	tctgttacc	agtggctgct	7980
gccagtggcg	ataagtctgt	tcttaccggg	ttggactcaa	gacgatagtt	accggataag	8040
gcgcagcgg	cgggctgaac	gggggggttc	tgcacacagc	ccagcttgg	gcgaacgacc	8100
tacaccgaac	tgagatacct	acagcgtgag	ctatgagaaa	gcgccacgct	tcccgagg	8160
agaaaggcgg	acaggtatcc	ggtaagcggc	agggtcggaa	caggagagcg	cacgaggag	8220
cttccagggg	gaaacgcctg	gtatctttat	agtcctgtcg	ggtttcgcca	cctctgactt	8280
gagcgtcgat	ttttgtgatg	ctcgtcaggg	ggcgaggacc	tatggaaaaa	cgccagcaac	8340
ggatgcgccg	cgtgcggctg	ctggagatgg	cggacgcgat	ggatatgttc	tgccaagggt	8400
tggtttg	attcacagtt	ctccgcaaga	attgattggc	tccaattctt	ggagtgggtga	8460
atccgttagc	gaggtgccgc	cggcttccat	tcaggtcgag	gtggcccggc	tccatgcacc	8520
gcgacgcaac	gcggggaggc	agacaaggta	tagggcgggc	cctacaatcc	atgccaaccc	8580
gttccatgtg	ctcgccgagg	cggcataaat	ccccgtgacg	atcagcggtc	caatgatcga	8640
agttaggctg	gtaagagccg	cgagcgatcc	ttgaagctgt	ccctgatgg	cgtcactctac	8700
ctgcctggac	agcatggcct	gcaacgcggg	catcccgatg	ccgccggaag	cgagaagaat	8760

36/36

cataatgggg	aaggccatcc	agcctcgcg	cgaggagctt	tttgcaaaag	cctaggcctc	8820
caaaaaagcc	tcctcactac	ttctggaata	gctcagaggc	cgaggcggcc	tcggcctctg	8880
cataaataaa	aaaaattagt	cagccatg				8908